

Training Programs for Improving Speech Perception in Noise: A Review

Nasrin Gohari¹, Zahra Hosseini Dastgerdi², Nematollah Rouhbakhsh³, Sara Afshar¹, and Razieh Mobini¹

¹Hearing Disorders Research Center, Department of Audiology, School of Rehabilitation, Hamadan University of Medical Sciences, Hamadan, Iran

²Department of Audiology, School of Rehabilitation, Isfahan University of Medical Sciences, Isfahan, Iran

³Department of Audiology, School of Rehabilitation, Tehran University of Medical Sciences, Tehran, Iran

Received July 3, 2022
Revised October 26, 2022
Accepted October 26, 2022

Address for correspondence

Zahra Hosseini Dastgerdi, PhD
Department of Audiology,
School of Rehabilitation,
Isfahan University of
Medical Sciences,
Isfahan, Iran
Tel +98-09132947800
Fax +98-(311)5145-668
E-mail zahra.au46@yahoo.com

Understanding speech in the presence of noise is difficult and challenging, even for people with normal hearing. Accurate pitch perception, coding and decoding of temporal and intensity cues, and cognitive factors are involved in speech perception in noise (SPIN); disruption in any of these can be a barrier to SPIN. Because the physiological representations of sounds can be corrected by exercises, training methods for any impairment can be used to improve speech perception. This study describes the various types of bottom-up training methods: pitch training based on fundamental frequency (F0) and harmonics; spatial, temporal, and phoneme training; and top-down training methods, such as cognitive training of functional memory. This study also discusses music training that affects both bottom-up and top-down components and speech training in noise. Given the effectiveness of all these training methods, we recommend identifying the defects underlying SPIN disorders and selecting the best training approach.

J Audiol Otol 2023;27(1):1-9

Keywords: Speech in noise perception; Speech in noise training; Rehabilitation.

Introduction

Accurate speech perception in everyday life is dependent on the hearing system's ability to process complex sounds in the presence of background noise. Speech perception in noise (SPIN) is also difficult for children and even some young adults with normal hearing and cognitive abilities. To understand speech in noisy situation, children with normal hearing require a higher signal-to-noise ratio (SNR) than adults. Because the most childhood learning takes place in noisy environments, the emergence of any speech perception disorder in children can result in learning, academic, and communication problems [1]. Adults with difficulty speech understanding in noise, on the other hand, complain of being tired of listening to or hearing something without understanding the meaning, discomfort from

background noise, and misunderstanding of conversations in the presence of competing sounds [2-4]. Speech impairment in noisy environments is one of the most important challenges for children with central auditory processing disorder (CAPD) [5], learning disabilities (LD) [6] attention deficit hyperactivity disorder (ADHD) [7], and hearing loss [8] and likewise for the elderly (over 65 years) [9-12]. The neurological and processing mechanisms associated with SPIN include pitch perception, neural coding and decoding of temporal and intensity cues, and cognitive skills. Each of these mechanisms is important for SPIN [13].

Pitch perception is an important indicator in the processing of complex stimuli and SPIN. Speech perception is related to the speaker's fundamental frequency (F0). The frequency of speech components is used to identify the speaker [14]. Studies have shown that listeners recognize F0 in the presence of ambient background noise and use other cues such as harmonics and formants [15]. Oxenham [16] showed that the ability to receive F0 and pitch is involved in concurrently grouping

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

and segregating speech sounds in normal hearing and people who have hearing loss or those using cochlear implants.

Temporal and intensity cues are related to interaural temporal difference (ITD) and interaural level difference (ILD), respectively. This is the primary condition for distinguishing target auditory data from non-target auditory data and also for normal auditory function [17,18]. Studies have shown that correct localization can help people with normal hearing thresholds to understand conversations with a lower SNR, and also demonstrated that auditory localization improves the SNR by 2 to 3 dB when the nature of the noise and signal are different, but increases the spatial advantage by 10 dB when the noise and signal are homogeneous [19,20]. When the noise and the signal have different frequency textures, the noise causes energetic masking. By contrast, when the noise and the signal of the same type coexist, e.g., target speech in the presence of distorted speech signals, the distracted noise causes both energetic and informational masking. Target signal processing in the presence of masking noises, particularly informational masking, necessitates the proper operation of cognitive mechanisms [21].

The main cognitive functions are attention, short-term memory, and working memory [3]. In cases such as hearing loss and the lack of temporal spectral encoding or sensory input, cognitive functions serve as compensatory mechanisms for the auditory system [22]. Speech perception exhibits a relationship with attention and auditory memory in the presence of noise [23,24]. Studies on patients with auditory attention and memory deficits showed that they had difficulty understanding speech in noise [25,26]. It can be hence concluded that higher levels of cognitive performance through top-down pathways can enhance bottom-up pathways and increase signal quality [3,27]. In addition, some exercises can influence both top-down and bottom-up activities [8].

The modification of the hearing system using educational tasks to strengthen speech comprehension is one of the most important and common areas of research today. Ample evidence suggests that hearing training exercises can help normal hearing and peoples with hearing loss, to improve spectral and temporal properties perception. Auditory training tasks are divided into three main categories: bottom-up, top-down, and mixed exercises (a combination of bottom-up and top-down methods). They aim to ameliorate auditory events comprehension via repetitive listening tasks [28]. Bottom-up tasks focus on the acoustic cues of the signals (i.e., spectral, temporal, and intensity characteristics), and require acoustic identification and differentiation. On the other side, top-down tasks improves signal perception by increasing attention to stimuli and encouraging the use of background information [28].

Various exercises and methods (bottom-up and top-down)

have been designed to improve SPIN in different clinical populations, including patients with auditory processing disorder (APD) [29], the elderly [30-32], hearing-impaired people [33], and autism patients [34]. This review study aims to search and analyze all the studies that investigated bottom-up and top-down hearing training programs for improving SPIN in different populations.

Training Programs Based on Cues and Factors Involved in Bottom-up Processes for Improving SPIN

Pitch training

Equivalent to the human sensory perception of sound frequency, the pitch is one of the psychological components of sound, along with loudness and timber, that aid in understanding music, speech perception, and sound separation in the presence of competing sound sources. Understanding the pitch of compound sounds depends on the discovery of the F0 and its harmonics and the periodicity of sounds [35]. Harmonics, along with other sound properties (spectral properties, fundamental frequency, synchronicity, or asynchronous at stimulus onset and offset), contribute to the formation of a hearing object by grouping/separating consecutive and concurrent sounds [36]. Because pitch is an important component of SPIN, various studies have investigated the training programs for improving pitch perception, including fundamental frequency [37,38] and harmonics trainings [8].

Fundamental frequency (F0) training

Changing the F0 is one of the methods examined in various studies on speech materials to improve pitch differentiation. Vowel differentiation has been used in F0 training studies. Vowels are components of verbal signals that contain a lot of information about F0, low-frequency components (first formant), and second formant. The acoustic information obtained from these formant features carries phonological information and, along with consonants, is considered a basis for speech comprehension [39]. The effect of F0 differences on the separation of concurrent sound sources is studied using concurrent vowel pairs [37,40]. These results show that increasing F0 differentiation between vowel and target speech from competitive speech improves speech perception [38].

Vowel interventions aim to improve the ability to distinguish between vowels. Meaningless syllables are frequently used in this training method to reduce the effects of top-down mechanisms (based on the schema and prior knowledge) and enhance vowel comprehension, as the basis of speech transmission and perception, by nourishing bottom-up pathways more

accurately. Vowels in various formats are presented during vowel training [41]. A review of the literature revealed that two studies have specifically investigated the vowel training interventions on hearing-impaired children [34] and the elderly with normal hearing [42]. A training pattern proposed by Talebi, et al. [33] involved teaching five vowel sounds (/æ/, /e/, /ɒ:/, /i:/, and /u:/) in a meaningless pattern of monosyllabic vowels including unvoiced syllables as /pæ/, /shæ/, /sæ/, /hæ/, and /kæ/. This training method is repeated for other sounds. The syllables were spoken behind the hearing impaired children, and they were asked to identify and produce syllables verbally. At the end of the training, the effect of this type of training on double-vowel separation was evaluated both behaviorally and electrophysiologically, and the results showed the effects of vowel training on concurrent sound separation [32,33]. There was a difference between pairs of vowels in their fundamental frequency [43]. Heidari, et al. [42] conducted a vowel training program based on hearing exercises for the elderly. The training program involved teaching six vowels (/æ/, /e/, /a/, /i/, /o/ and /u/) as meaningless monosyllabic vowels for a period of 5 weeks. For example, the monosyllabic vowels /pæ/, /fæ/, /sæ/, /hæ/, and /kæ/ were spoken at a distance of one meter behind the person at a hearing comfortable level. This process was repeated for the vowels /e/, /a/, /i/, /o/ and /u/. This training program also improved SPIN among the elderly.

Harmonic training

Mistuning harmonics is a behavioral [8] and electrophysiological [44] method of evaluating concurrent sound separation. It is assumed that when one of the harmonic components is mistuned by 3% of its harmonic value, it is heard as a separate tone from the complex tone. Mistuned harmonics produce and perceive beats as a result of amplitude modulation in the temporal envelope of the stimulus waveform. Increasing the mistuning rate from 3% to 16% and the duration to more than 50 milliseconds results in a decrease in the detection threshold of beats [36]. Moossavi, et al. [8] tested the ability of several hearing-impaired children to perceive and distinguish harmonics and, for the first time, designed a new method of harmonic differentiation training. According to the prevalence of the fundamental frequency of human speech sounds in men (100–146 Hz) and women (188–221 Hz) [45], they employed a complex tone of 100, 200, and 300 Hz with their first 10 harmonics and mistuned one of the first harmonics each time by 2%, 4%, 8%, and 16%. Then they asked the children to compare the complex tone with the tuned harmonic and the mistuned harmonic. The training was performed at frequencies of 100, 200, and 300 Hz in the first, third, fifth, seventh, and ninth harmonics after obtaining the mistuning

differentiation threshold in each harmonic of the complex tone. The training began with the lowest level of detection in each harmonic and increased by 70% of the difficulty level if differentiated. The study findings revealed that this training improved the performance of hearing-impaired children in the frequency pattern sequence test and consonant-vowel perception in noise [8].

Localization training

For sound localization, the human auditory system employs two delicate cues. Horizontal level localization is based on ITD for sounds below 1,500 Hz, and ILD and spectral cues for sounds above 2,500 Hz. These spatial and spectral cues may be important for spatial stream segregation, which aids in speech perception [46]. ITD is more important to sound localization [47]. The ITD in unmodulated signals is only processed up to 1,500 Hz and is referred to as the ITD fine structure (ITD FS) [48,49]. A slow carrier (low frequency) modulates the ITD information at higher frequencies, and this is known as ITD envelope (ITD ENV) [50]. Speech, as a modulated signal, contains two types of ITD: ITD FS and ITD ENV [51].

Wright and Fitzgerald [52] evaluated the effectiveness of training based on the simulation of ITD and ILD cues under headphones and reported the improved ability of participants to diagnose ITD and ILD after training. Furthermore, a study by Kuk, et al. [53] demonstrated the reliability of the localization training program in people with hearing loss. Sound Storm (previously known as the Listen & Learn Auditory Training Software) is a training program designed and evaluated by Cameron, et al. [54] to improve the binaural processing skills of children with spatial processing disorder (SPD). SPD is defined as the inability of individuals to use binaural cues to selectively pay attention to sounds coming from one direction while suppressing sounds coming from other directions. As a result, children with SPD struggle to understand speech in noisy environments, such as classrooms. The Sound Storm software creates a three-dimensional hearing environment through headphones. The child is required to repeat a word, including the target, which appears in the background noise. A comparative method according to the patient's response determines the level of sentence intensity. In their study, they used the Sound Storm auditory training software to train 9 suspected SPD children who performed abnormally in the speech-in-noise assessment (LiSN-S PGA) for 15 minutes per day for a period of 12 weeks. They examined the effect of the intervention after three months and found that the speech perception threshold improved by 10 dB in the Sound Storm auditory training software [54].

Lotfi, et al. [29] studied the effectiveness of an auditory

localization rehabilitation program based on ITD cues on spatialized speech in noise and monaural low redundancy tests in a group of children with suspected APD. After 12 training sessions, the mean speech perception score was significantly better in the experimental group in comparison to the control group. Lotfi, et al. [30] also investigated the effect of a spatial processing training program called the Persian spatialized speech in noise test on speech perception skills in the elderly with normal hearing who complaint of difficulty with SPIN. Their findings revealed that the SNR significantly reduced to 50% in speech perception after the training, whereas auditory spatial rehabilitation could assist the older adults in benefiting from the spatial difference of sound sources and noise for better speech perception.

Recent investigations have dealt with the role of ITD ENV in spatial hearing and SPIN [55,56]. Majdek, et al. [51] confirmed the role of ITD ENV in speech localization and understanding in noisy environments. In a study by Delphi, et al. [31] stimulus ENV in various ITDs were presented to the elderly with normal hearing. The ENV was designed in 10-millisecond steps from 10 to 100 milliseconds and 50-millisecond steps from 100 to 350 milliseconds. ITD ENV was initially implemented according to the results of previous assessments and then it was gradually reduced. Each ITD ENV stimulus was presented again and again until the participant correctly identified it. SPIN was re-evaluated at the end of the training to measure the effectiveness of the training. The results revealed that the ITD ENV training not only improved the localization abilities of participants but also increased their mean score of SPIN.

Temporal training

The temporal processing is a broad category of auditory processing abilities that includes temporal sequencing, temporal integration, temporal resolution, and temporal masking [32]. According to perceptual and neurophysiological studies, temporal processing skills are more affected by age than other central auditory processing skills. There are numerous biological reasons for the auditory system's poor functioning in the elderly, including decreased myelin integrity, nerve long-term recovery, decreased brain connections, and decreased neural synchronicity [57]. Inability to process time can be a major cause of language disorders. For example, autism has been linked to temporal processing disorders and their relationship with SPIN [34]. Because of the importance of time processing, it is necessary to develop training approaches based on this skill. Temporal training programs in the form of formal and organized homework, and even computer games, have been used in studies with various goals [58]. One of the most

important goals of temporal processing training is to improve time and language processing skills, as well as speech perception in noisy environments.

Ramezani, et al. [34] assessed the effectiveness of temporal processing based rehabilitation on sequencing skills and temporal resolution using temporal pattern detection and noise detection exercises on the speech perception in noise ability and on the speech ABR components in 28 autistic adolescents. Their findings showed a significant improvement in SPIN and the efficiency of time processing at the auditory brainstem level (reduction of wave latency) in response to speech signals. Sattari, et al. [32] also designed the rehabilitation tasks related to auditory temporal processing in 5 domains, including detection of the number of stimuli, pitch detection, detection of duration patterns, detection of the number of meaningless speech stimuli in noise, and gap detection in noise, to improve SPIN in the elderly aged 75–60 years who were using hearing aids. In another study, Rasouli Fard, et al. [59] separated the fine temporal structure and the stimuli envelope of vowel-consonant-vowel (VCV) and then designed a training method based on the fine temporal structure and differentiation of 16 consonants. Their results showed an improvement in SPIN among the elderly with mild to moderate hearing loss.

Phonemic training

Bottom-up processes (e.g., as sensory speech processing) and top-down processes (e.g., selective attention, short- and long-term memory, and the use of lexical and textual concepts) are always involved in the process of speech perception. Both types of processes are constantly interacting with one another and are inseparable. By reducing semantic information (content and context), individuals rely on phonetic features (such as formant frequencies and voice onset time) for speech perception [60]. However, it has been reported that bottom-up training protocols improve everyday listening ability less than top-down training [61]. Schumann, et al. [60] reported that a computer-based listening training program, which included phonological differentiation using VCV and consonant-vowel-consonant (CVC), was effective in improving SPIN in individuals using cochlear implants. As the most common disorder in auditory processing, phonological differentiation training is one of the components of defect decoding training in the Buffalo model of therapy for APDs [62]. This training program, which is known as phonemic training program, aims to improve speech comprehension, reading ability, auditory spelling improvement, and speech clarity. According to the results of phonemic error analysis, this training program begins with easier tasks, such as providing phonemes that are easier to perceive for participants, and gradually increases in difficulty as auditory processing goes

on. Practices in this training program are based on the repetition and differentiation of learned phonemes [39].

Training Methods for Improving SPIN Based on Factors Involved in Top-Down Processes

Memory-based training

Many studies have extensively dealt with cognitive training in recent years and demonstrated its remarkable effects on improving academic performance [63], and language learning [64], and lowering the risk of dementia [65]. Cognitive training includes attention, functional memory, and short-term memory reinforcement [3]. However, more emphasis is usually placed on improving functional memory [63,66]. As part of the cognitive process, memory related to the reception, process and store of verbal stimuli, and eventually recall what have been heard. Auditory memory serves as the foundation for the development of language skills (including the ability to learn and memorize words as well as the ability to understand and apply grammar, spoken language, and written language) and the learning process. As a result, language could not be imagine without memory [67]. More extensive functional memory capacities are associated with improved academic performance and language learning as well as a lower risk of pathological aging [63,66]. After functional memory training, the elderly with unknown cognitive impairment showed an increase in functional memory capacity [68,69]. Functional memory has also been shown to play a role in listeners' ability to understand speech in noisy environments. Among people with normal hearing and those with hearing impairments, listeners with greater functional memory capacity appear to be more successful in understanding speech in noise. Greater functional memory capacity predicts greater success in speech recognition in noise in older people who use hearing aids [70,71]. Because functional memory capacity predicts the success of SPIN, it can be suggested that cognitive training and increasing listeners' functional memory capacity can be a basis for developing an effective intervention. Many functional auditory memory training tasks rely on straightforward test items including numbers, letters, or monosyllabic words. In addition, since these assignments do not have a high semantic load, they are relatively strong interventions for improving hearing loss [72]. Ingvalson, et al. [73] reported that 10 days of reverse digit training significantly improved reading comprehension and SPIN among native Mandarin Chinese speakers and native Chinese English speakers. Therefore, this type of training can be used to improve SPIN.

Music training

Music is a source of pleasure, learning, and well-being. Prolonged exposure to music creates plasticity from the cochlea to the auditory cortex [74]. Several studies have shown that exposure to music can lead to the transfer of specific listening proficiencies to non-musical domains [74,75]. The theories presented by Patel shows that music education increases adaptive flexibility in speech processing neural networks [76]. It also reinforces numerous auditory processing components involved SPIN among musicians, including syllable differentiation [77,78], speech temporal signals processing [78], prosody [79], pitch and harmonics [80], melodic contour [81], and auditory cognition such as working memory [82,83], attention [84,85], and neural representation of speech [86,87].

Zendel and Alain [88] employed ERP responses to investigate the separation of concurrent sounds in musicians with a mean age of 28 years. N1 and P2 waves were used as evidence to study the music training effectiveness on the separation process. The findings demonstrated the musicians' exceptional ability to separate concurrent sounds. The researchers in this study hypothesized that training methods, such as music therapy, could have a significant effect on the separation of sounds presented concurrently, ultimately improving sound perception (especially speech). In another study conducted by Swaminathan, et al. [89], the spatial release of noise was higher among musicians than non-musicians. Parbery-Clark, et al. [90] demonstrated improvement in speech perception in both multi-talker and continuous pseudo-speech noise in musicians.

Considering the benefits of music in improving the performance of bottom-up and top-down auditory processing as well as the better performance of musicians in SPIN, Slater, et al. [91] used a long-term music training (2 years) to improve children's ability to understand speech in noisy environments. Jain, et al. [92] performed an 8-day music training for adults aged 18–25 years and observed an improvement in SPIN. It has also been shown that short-term music training in the elderly improves neural speech coding and facilitates SPIN [93]. Jayakody [94] examined the effects of a computer-based music perception training program and concluded that this program improved music comprehension and speech perception in people using cochlear implants or hearing aids, more significantly in those who were using hearing aids.

Speech-in-noise training

Speech-in-noise training (SINT), which focuses on improving listening skills in noisy environments, is an important component of an auditory training program for the hearing impaired and a wide range of individuals with hearing impairment, learning disabilities, and LD [95]. Various speech materials, includ-

ing syllables, words (monosyllabic and bisyllabic), sentences, phrases, and texts, at specific SNRs can be used in SINT. Such methods, which were first described by Katz and Burge [96] employed noise desensitization techniques to improve SPIN. This method increases the level of noise tolerance and memory and, as a result, improves SPIN. This technique helps participants desensitize to noise and develop strategies to obtain more information about auditory signals in difficult-to-hear situations [56]. This training program involves listening to speech stimuli at different levels of noise, because exposure to a gradually increasing level of noise reduces the effect of noise. In this program, the words are first presented in silence and then in noise, and the SNR is gradually increased from easy (e.g., +12 or +15 dB) to difficult (0 dB or negative SNRs) conditions [97]. Masters propose that a treatment hierarchy begins with white noise and ends up with the most problematic type of noise for the client. It is also suggested that such interventions take into account the type of noises that one is more exposed to in the surrounding environment [98]. Several studies have reported the effectiveness of noise speech training, particularly the noise desensitization method. According to Jutras, et al. [99], children with APD exhibited a significant increase in noise tolerance after receiving noise-specific speech training. They investigated how SINT affected speech perception test scores, electrophysiological criteria, and listening and living habits in the child's real-life conditions and environment. Speech perception test scores and electrophysiological components improved due to this rehabilitation. Furthermore, improvement was not limited to noisy environments but it happened in other challenging situations of everyday life [99]. Kumar, et al. [95] designed a computerized tasks for speech in noise training and evaluated its effectiveness using auditory processing behavioral tests and auditory electrophysiological responses. This computer-based program (module) included teaching words in noise, in which single and three-syllable words were presented as target words in the presence of spoken noise and multi-talker babble at various SNRs ranging from +20 to -4 dB. The performance of experimental children in speech noise and auditory processing was better than the control group. In addition, after training, the range of evoked responses with high latency in silence and noise decreased significantly. Their study found that a targeted program of SINT improved auditory behavioral skills and electrophysiological responses. Another study in 2021 investigated the effects of a noise word training program using monosyllabic words in the presence of spoken and multi-speaker noise in the SNR range of +20 to 4 dB on processing and cognitive skills (working memory) of 20 children with APD (10 children in the experimental group and 10 children in the control group). The results of this study

showed that SINT significantly improved the mean SPIN score, temporal processing (gap in noise test and duration pattern), and the backward digit's span score [100].

To design speech tasks in noise, it is recommended that noise be used in conjunction with the speech signal in a systematic manner, beginning with energetic noise and progressing to informational noise. Energetic masking is associated with noise that has a high-frequency compression and no gap between the energy spectrums produced, such as white noise. In this case, it seems that the neural networks in charge of signal and noise processing are sensitive to the all energy level of the signal and noise, and thus the signal with a slightly higher absolute energy level than the noise is easily detected. Speech comprehension in a noisy environment will be possible with lower SNRs and the use of pitch and spatial cues. The difficulty and the challenging level of the auditory environment are controlled and adjusted in this type of test based on SNR. Providing speech in the presence of informational masking, on the other hand, targets a different level of ability of the neural networks responsible for signal processing in noise. Distracting noises of the signal type are referred to as informational masking. When speech is in the speech noise, a person's ability to understand usually suffers. It is preferable to begin with multiple speech streams and end up with two speech streams (from easy to hard). In the presence of informational noise, the auditory system typically uses signal monitoring in the ears at any frequency and at any time, according to theories of speech signal detection. In other words, it is assumed that our auditory system searches for target signal among the noise gaps or tries to collect the target signal from the better ear with higher SNR, an approach which is known as cross-ear dip-listening [101-104].

Conclusion

Difficulty with SPIN is a common condition affecting all children and adults with central hearing impairment, children with LD, ADHD, autism, and hearing loss, and the elderly. Each person with this problem suffers from a defect in one or more of the underlying neural mechanisms of SPIN, which can be improved through appropriate training programs. Pitch training, spatial training, temporal training, phoneme training, functional memory training, musical training, and SINT are all used to improve SPIN. Pitch training is based on fundamental frequency and harmonics. Changing the fundamental frequency is one of the methods to improve pitch differentiation based on vowel differentiation at various fundamental frequencies. Mistuning complex tone harmonics and distinguishing tuning stimuli from mistuning from the base of harmonic training. Localization train-

ing involves simulating the space environment with headphones and investigating localization disorders in various ITDs. The detection of the number of stimuli, pitch stimuli, duration pattern, number of meaningless speech stimuli in noise, and the gap in noise is the central part of temporal training, whereas phoneme training focuses on meaningless syllables of VCV and CVC. Memory training is performed for patients with poor memory using simple stimuli, such as numbers and letters. Music training has been shown to have a significant effect on SPIN by improving the reception of acoustic stimuli and involving top-down processes of attention and memory. Finally, SINT using stimuli, such as syllables, words, and sentences, in energetic and informational background noise has been shown to improve SPIN.

Conflicts of Interest

The authors have no financial conflicts of interest.

Author Contributions

Conceptualization: Nasrin Gohari, Zahra Hosseini Dastgerdi. Data curation: Nasrin Gohari, Zahra Hosseini Dastgerdi. Formal analysis: Nasrin Gohari, Zahra Hosseini Dastgerdi. Investigation: all authors. Methodology: Nasrin Gohari, Zahra Hosseini Dastgerdi, Nematollah Rouhbakhsh. Project administration: Nasrin Gohari, Zahra Hosseini Dastgerdi. Supervision: Nasrin Gohari, Zahra Hosseini Dastgerdi. Validation: Nasrin Gohari, Nematollah Rouhbakhsh, Sara Afshar, Razieh Mobini. Visualization: Nasrin Gohari, Nematollah Rouhbakhsh, Sara Afshar, Razieh Mobini. Writing—original draft: all authors. Writing—review & editing: Nasrin Gohari, Zahra Hosseini Dastgerdi. Approval of final manuscript: all authors.

ORCID iDs

Nasrin Gohari <https://orcid.org/0000-0002-3519-5878>
 Zahra Hosseini Dastgerdi <https://orcid.org/0000-0002-6376-6033>
 Nematollah Rouhbakhsh <https://orcid.org/0000-0003-2670-0139>
 Sara Afshar <https://orcid.org/0000-0003-2407-0938>
 Razieh Mobini <https://orcid.org/0000-0003-2249-0701>

REFERENCES

- Shield BM, Dockrell JE. The effects of noise on children at school: a review. *Build Acoust* 2003;10:97-116.
- Gordon-Salant S, Yeni-Komshian GH, Fitzgibbons PJ, Barrett J. Age-related differences in identification and discrimination of temporal cues in speech segments. *J Acoust Soc Am* 2006;119:2455-66.
- Strait DL, Parbery-Clark A, O'Connell S, Kraus N. Biological impact of preschool music classes on processing speech in noise. *Dev Cogn Neurosci* 2013;6:51-60.
- Bradlow AR, Kraus N, Hayes E. Speaking clearly for children with learning disabilities: sentence perception in noise. *J Speech Lang Hear Res* 2003;46:80-97.
- Flanagan S, Zorilá TC, Stylianou Y, Moore BCJ. Speech processing to improve the perception of speech in background noise for children with auditory processing disorder and typically developing peers. *Trends Hear* 2018;22:2331216518756533.
- Ferency M, Pottas L, Soer M. Speech perception in noise in children with learning difficulties: a scoping review. *Int J Pediatr Otorhinolaryngol* 2022;156:111101.
- Blomberg R, Danielsson H, Rudner M, Söderlund GBW, Rönnerberg J. Speech processing difficulties in attention deficit hyperactivity disorder. *Front Psychol* 2019;10:1536.
- Moossavi A, Mehrkian S, Gohari N, Nazari MA, Bakhshi E, Alain C. The effect of harmonic training on speech perception in noise in hearing-impaired children. *Int J Pediatr Otorhinolaryngol* 2021;149:110845.
- Warrier CM, Johnson KL, Hayes EA, Nicol T, Kraus N. Learning impaired children exhibit timing deficits and training-related improvements in auditory cortical responses to speech in noise. *Exp Brain Res* 2004;157:431-41.
- Anderson S, Skoe E, Chandrasekaran B, Kraus N. Neural timing is linked to speech perception in noise. *J Neurosci* 2010;30:4922-6.
- Souza PE, Boike KT, Witherell K, Tremblay K. Prediction of speech recognition from audibility in older listeners with hearing loss: effects of age, amplification, and background noise. *J Am Acad Audiol* 2007;18:54-65.
- Zendel BR, West GL, Belleville S, Peretz I. Musical training improves the ability to understand speech-in-noise in older adults. *Neurobiol Aging* 2019;81:102-15.
- Aarabi S, Jarollahi F, Badfar S, Hosseinabadi R, Ahadi M. Speech perception in noise mechanisms. *Aud Vest Res* 2016;25:221-6.
- Stickney GS, Assmann PF, Chang J, Zeng FG. Effects of cochlear implant processing and fundamental frequency on the intelligibility of competing sentences. *J Acoust Soc Am* 2007;122:1069-78.
- Baumann O, Belin P. Perceptual scaling of voice identity: common dimensions for different vowels and speakers. *Psychol Res* 2010;74:110-20.
- Oxenham AJ. Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants. *Trends Amplif* 2008;12:316-31.
- King AJ, Dahmen JC, Keating P, Leach ND, Nodal FR, Bajo VM. Neural circuits underlying adaptation and learning in the perception of auditory space. *Neurosci Biobehav Rev* 2011;35:2129-39.
- Vickers NJ. Animal communication: when I'm calling you, will you answer too? *Curr Biol* 2017;27:R713-5.
- Noble W, Gatehouse S. Effects of bilateral versus unilateral hearing aid fitting on abilities measured by the speech, spatial, and qualities of hearing scale (SSQ). *Int J Audiol* 2006;45:172-81.
- Ramsden JD, Papsin BC, Leung R, James A, Gordon KA. Bilateral simultaneous cochlear implantation in children: our first 50 cases. *Laryngoscope* 2009;119:2444-8.
- McCreery RW, Walker EA, Spratford M, Lewis D, Brennan M. Auditory, cognitive, and linguistic factors predict speech recognition in adverse listening conditions for children with hearing loss. *Front Neurosci* 2019;13:1093.
- Wong PC, Jin JX, Gunasekera GM, Abel R, Lee ER, Dhar S. Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia* 2009;47:693-703.
- Strait DL, Kraus N. Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Front Psychol* 2011;2:113.
- Obleser J, Wise RJ, Dresner MA, Scott SK. Functional integration across brain regions improves speech perception under adverse listening conditions. *J Neurosci* 2007;27:2283-9.
- Tun PA, O'Kane G, Wingfield A. Distraction by competing speech in young and older adult listeners. *Psychol Aging* 2002;17:453-67.
- Frisina DR, Frisina RD. Speech recognition in noise and presbycusis: relations to possible neural mechanisms. *Hear Res* 1997;106:95-104.
- Edeline JM. The thalamo-cortical auditory receptive fields: regulation by the states of vigilance, learning and the neuromodulatory systems. *Exp Brain Res* 2003;153:554-72.
- Alain C, Snyder JS, He Y, Reinke KS. Changes in auditory cortex parallel rapid perceptual learning. *Cereb Cortex* 2007;17:1074-84.
- Lotfi Y, Moosavi A, Abdollahi FZ, Bakhshi E, Sadjedi H. Effects of an auditory lateralization training in children suspected to cen-

- tral auditory processing disorder. *J Audiol Otol* 2016;20:102-8.
- 30) Lotfi Y, Samadi-Qaleh-Juqy Z, Moosavi A, Sadjedi H, Bakhshi E. The effects of spatial auditory training on speech perception in noise in the elderly. *Crescent J Med Biol Sci* 2020;7:40-6.
 - 31) Delphi M, Lotfi MY, Moosavi A, Bakhshi E, Banimostafa M. Reliability of interaural time difference-based localization training in elderly individuals with speech-in-noise perception disorder. *Iran J Med Sci* 2017;42:437-42.
 - 32) Sattari K, Rahbar N, Ahadi M, Haghani H. The effects of a temporal processing-based auditory training program on the auditory skills of elderly users of hearing aids: a study protocol for a randomized clinical trial. *F1000Res* 2020;9:425.
 - 33) Talebi H, Moosavi A, Lotfi Y, Faghihzadeh S. Effects of vowel auditory training on concurrent speech segregation in hearing impaired children. *Ann Otol Rhinol Laryngol* 2014;124:13-20.
 - 34) Ramezani M, Lotfi Y, Moosavi A, Bakhshi E. Effects of auditory processing training on speech perception and brainstem plasticity in adolescents with autism spectrum disorders. *Iran J Child Neurol* 2021;15:69-77.
 - 35) Graves JE, Oxenham AJ. Pitch discrimination with mixtures of three concurrent harmonic complexes. *J Acoust Soc Am* 2019;145:2072-83.
 - 36) Moore DR, Fuchs PA, Rees A, Palmer A, Plack CJ. *The Oxford handbook of auditory science: the auditory brain*. Oxford: Oxford University Press;2010.
 - 37) de Cheveigné A, Kawahara H, Tsuzaki M, Aikawa K. Concurrent vowel identification. I. Effects of relative amplitude and F0 difference. *J Acoust Soc Am* 1997;101:2839-47.
 - 38) Broxk JPL, Nooteboom SG. Intonation and the perceptual separation of simultaneous voices. *J Phon* 1982;10:23-36.
 - 39) Katz J, Chasin M, English KM, Hood LJ, Tillery KL. *Handbook of clinical audiology*. Philadelphia, PA: Wolters Kluwer Health;2015.
 - 40) Assmann PF, Summerfield Q. Modeling the perception of concurrent vowels: vowels with different fundamental frequencies. *J Acoust Soc Am* 1990;88:680-97.
 - 41) Tye-Murray N. *Foundations of aural rehabilitation: children, adults, and their family members*. San Diego, CA: Plural Publishing;2019.
 - 42) Heidari A, Moosavi A, Yadegari F, Bakhshi E, Ahadi M. Effect of vowel auditory training on the speech-in-noise perception among older adults with normal hearing. *Iran J Otorhinolaryngol* 2020;32:229-36.
 - 43) Snyder JS, Alain C. Age-related changes in neural activity associated with concurrent vowel segregation. *Brain Res Cogn Brain Res* 2005;24:492-9.
 - 44) Mehrkian S, Moosavi A, Gohari N, Nazari MA, Bakhshi E, Alain C. Long latency auditory evoked potentials and object-related negativity based on harmonicity in hearing-impaired children. *Neurosci Res* 2022;178:52-9.
 - 45) Baken RJ, Orlikoff RF. *Clinical measurement of speech and voice*. Boston: Cengage Learning;2000.
 - 46) Van Deun L, van Wieringen A, Van den Bogaert T, Scherf F, Offeciers FE, Van de Heyning PH, et al. Sound localization, sound lateralization, and binaural masking level differences in young children with normal hearing. *Ear Hear* 2009;30:178-90.
 - 47) Babkoff H, Muchnik C, Ben-David N, Furst M, Even-Zohar S, Hildesheimer M. Mapping lateralization of click trains in younger and older populations. *Hear Res* 2002;165:117-27.
 - 48) Gilkey R, Anderson TR. *Binaural and spatial hearing in real and virtual environments*. New York: Psychology Press;2014.
 - 49) Brughera A, Dunai L, Hartmann WM. Human interaural time difference thresholds for sine tones: the high-frequency limit. *J Acoust Soc Am* 2013;133:2839-55.
 - 50) Bernstein LR. Auditory processing of interaural timing information: new insights. *J Neurosci Res* 2001;66:1035-46.
 - 51) Majdak P, Laback B, Baumgartner WD. Effects of interaural time differences in fine structure and envelope on lateral discrimination in electric hearing. *J Acoust Soc Am* 2006;120:2190-201.
 - 52) Wright BA, Fitzgerald MB. Different patterns of human discrimination learning for two interaural cues to sound-source location. *Proc Natl Acad Sci U S A* 2001;98:12307-12.
 - 53) Kuk F, Keenan DM, Lau C, Crose B, Schumacher J. Evaluation of a localization training program for hearing impaired listeners. *Ear Hear* 2014;35:652-66.
 - 54) Cameron S, Dillon H. Development and evaluation of the LiSN & learn auditory training software for deficit-specific remediation of binaural processing deficits in children: preliminary findings. *J Am Acad Audiol* 2011;22:678-96.
 - 55) Joris PX. Interaural time sensitivity dominated by cochlea-induced envelope patterns. *J Neurosci* 2003;23:6345-50.
 - 56) Laback B, Pok SM, Baumgartner WD, Deutsch WA, Schmid K. Sensitivity to interaural level and envelope time differences of two bilateral cochlear implant listeners using clinical sound processors. *Ear Hear* 2004;25:488-500.
 - 57) Fitzgibbons PJ, Gordon-Salant S. Age-related differences in discrimination of temporal intervals in accented tone sequences. *Hear Res* 2010;264:41-7.
 - 58) Weihing J, Chermak GD, Musiek FE. Auditory training for central auditory processing disorder. *Semin Hear* 2015;36:199-215.
 - 59) Rasouli Fard P, Jarollahi F, Sameni SJ, Kamali M. Development of a training software to improve speech-in-noise perception in the elderly with noise-induced hearing loss. *Aud Vestib Res* 2022;31:38-44.
 - 60) Schumann A, Serman M, Gefeller O, Hoppe U. Computer-based auditory phoneme discrimination training improves speech recognition in noise in experienced adult cochlear implant listeners. *Int J Audiol* 2015;54:190-8.
 - 61) Sweetow R, Palmer CV. Efficacy of individual auditory training in adults: a systematic review of the evidence. *J Am Acad Audiol* 2005;16:494-504.
 - 62) Lucker JR. Phonemic awareness, reading abilities, and auditory processing disorders. In: *Auditory Processing Disorders: Assessment, Management, and Treatment* (eds. Geffner D, Ross-Swain D). San Diego, CA: Plural Publishing;2018. p.391.
 - 63) Klingberg T. Training and plasticity of working memory. *Trends Cogn Sci* 2010;14:317-24.
 - 64) Inghvalson EM, Barr AM, Wong PC. Poorer phonetic perceivers show greater benefit in phonetic-phonological speech learning. *J Speech Lang Hear Res* 2013;56:1045-50.
 - 65) Willis SL, Tennstedt SL, Marsiske M, Ball K, Elias J, Koepke KM, et al. Long-term effects of cognitive training on everyday functional outcomes in older adults. *JAMA* 2006;296:2805-14.
 - 66) Morrison AB, Chein JM. Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychon Bull Rev* 2011;18:46-60.
 - 67) Cusimano A. *Learning disabilities: there is a cure*. Lansdale, PA: Achieve Publications, Inc;2002.
 - 68) Bherer L, Kramer AF, Peterson MS, Colcombe S, Erickson K, Bécic E. Training effects on dual-task performance: are there age-related differences in plasticity of attentional control? *Psychol Aging* 2005;20:695-709.
 - 69) Li SC, Schmiedek F, Huxhold O, Röcke C, Smith J, Lindenberger U. Working memory plasticity in old age: practice gain, transfer, and maintenance. *Psychol Aging* 2008;23:731-42.
 - 70) Foo C, Rudner M, Rönnerberg J, Lunner T. Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. *J Am Acad Audiol* 2007;18:618-31.
 - 71) Lunner T, Sundewall-Thorén E. Interactions between cognition, compression, and listening conditions: effects on speech-in-noise performance in a two-channel hearing aid. *J Am Acad Audiol* 2007;18:604-17.
 - 72) Smits C, Theo Goverts S, Festen JM. The digits-in-noise test: as-

- sessing auditory speech recognition abilities in noise. *J Acoust Soc Am* 2013;133:1693-706.
- 73) Ingvalson EM, Dhar S, Wong PC, Liu H. Working memory training to improve speech perception in noise across languages. *J Acoust Soc Am* 2015;137:3477-86.
 - 74) Moossavi A, Gohari N. The impact of music on auditory and speech processing. *Audit Vestib Res* 2019;28:134-45.
 - 75) Peretz I. Music, language, and modularity in action. In: *Language and Music as Cognitive Systems* (eds. Rebuschat P, Rohrmeier M, Hawkins JA, Cross I). Oxford: Oxford University Press;2012. p.254-68.
 - 76) Patel AD. Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Front Psychol* 2011;2:142.
 - 77) Zuk J, Ozernov-Palchik O, Kim H, Lakshminarayanan K, Gabrieli JD, Tallal P, et al. Enhanced syllable discrimination thresholds in musicians. *PLoS One* 2013;8:e80546.
 - 78) Parbery-Clark A, Tierney A, Strait DL, Kraus N. Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience* 2012; 219:111-9.
 - 79) Magne C, Schön D, Besson M. Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *J Cogn Neurosci* 2006;18:199-211.
 - 80) Micheyl C, Delhommeau K, Perrot X, Oxenham AJ. Influence of musical and psychoacoustical training on pitch discrimination. *Hear Res* 2006;219:36-47.
 - 81) Fujioka T, Trainor LJ, Ross B, Kakigi R, Pantev C. Musical training enhances automatic encoding of melodic contour and interval structure. *J Cogn Neurosci* 2004;16:1010-21.
 - 82) George EM, Coch D. Music training and working memory: an ERP study. *Neuropsychologia* 2011;49:1083-94.
 - 83) Schulze K, Zysset S, Mueller K, Friederici AD, Koelsch S. Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. *Hum Brain Mapp* 2011;32:771-83.
 - 84) Strait DL, Kraus N, Parbery-Clark A, Ashley R. Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. *Hear Res* 2010;261:22-9.
 - 85) Baumann S, Meyer M, Jäncke L. Enhancement of auditory-evoked potentials in musicians reflects an influence of expertise but not selective attention. *J Cogn Neurosci* 2008;20:2238-49.
 - 86) Bidelman GM, Krishnan A. Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Res* 2010;1355:112-25.
 - 87) Zendel BR, Tremblay CD, Belleville S, Peretz I. The impact of musicianship on the cortical mechanisms related to separating speech from background noise. *J Cogn Neurosci* 2015;27:1044-59.
 - 88) Zendel BR, Alain C. Concurrent sound segregation is enhanced in musicians. *J Cogn Neurosci* 2009;21:1488-98.
 - 89) Swaminathan J, Mason CR, Streeter TM, Kidd Jr G, Patel AD. Spatial release from masking in musicians and non-musicians. *J Acoust Soc Am* 2014;135:2281-2.
 - 90) Parbery-Clark A, Skoe E, Lam C, Kraus N. Musician enhancement for speech-in-noise. *Ear Hear* 2009;30:653-61.
 - 91) Slater J, Skoe E, Strait DL, O'Connell S, Thompson E, Kraus N. Music training improves speech-in-noise perception: longitudinal evidence from a community-based music program. *Behav Brain Res* 2015;291:244-52.
 - 92) Jain C, Mohamed H, Kumar AU. The effect of short-term musical training on speech perception in noise. *Audiol Res* 2015;5:111.
 - 93) Fleming D, Belleville S, Peretz I, West G, Zendel BR. The effects of short-term musical training on the neural processing of speech-in-noise in older adults. *Brain Cogn* 2019;136:103592.
 - 94) Jayakody DMP. A computerized pitch-perception training program for the hearing impaired [dissertation]. Christchurch: University of Canterbury;2011.
 - 95) Kumar P, Singh NK, Hussain RO. Efficacy of computer-based noise desensitization training in children with speech-in-noise deficits. *Am J Audiol* 2021;30:325-40.
 - 96) Katz J, Burge C. Auditory perception training for children with learning disabilities. *Menorah Medical Journal* 1971;2:18-29.
 - 97) Maggu AR, Yathiraj A. Effect of noise desensitization training on children with poor speech-in-noise scores. *Can J Speech-Lang Pathol Audiol* 2011;35:56-65.
 - 98) Westers MG, Stecker NA, Katz J. Central auditory processing disorders: mostly management. Needham Heights, MA: Allyn & Bacon;1998.
 - 99) Jutras B, Lafontaine L, East MP, Noël M. Listening in noise training in children with auditory processing disorder: exploring group and individual data. *Disability and Rehabilitation* 2019;41:2918-26.
 - 100) Kumar P, Singh NK, Hussain RO. Effect of speech in noise training in the auditory and cognitive skills in children with auditory processing disorders. *Int J Pediatr Otorhinolaryngol* 2021;146:110735.
 - 101) Buchholz JM, Dillon H, Cameron S. Towards a listening in spatialized noise test using complex tones. *Proc Mtgs Acoust* 2013;19: 050047.
 - 102) Westermann A, Buchholz JM. The influence of informational masking in reverberant, multi-talker environments. *J Acoust Soc Am* 2015;138:584-93.
 - 103) Westermann A. Understanding speech in complex acoustic environments: the role of informational masking and auditory distance perception [dissertation]. Macquarie Park: Macquarie Univ.;2015.
 - 104) Ozmeral EJ. The role of upward spread of masking in the ability to benefit from asynchronous glimpsing of masked speech [dissertation]. Chapel Hill, NC: University of North Carolina at Chapel Hill; 2011.